

**Quality Measures Workgroup:
Methodologic Tiger Team
Draft Transcript
October 25, 2010**

Presentation

Jon White – AHRQ/HHS – Director IT

Hello, everybody. This is Jon White from the Agency for Healthcare Research and Quality. I want to thank you for joining us for today's Health IT Policy Committee Quality Measures Workgroup Methodological Issues Tiger Team. It's a mouthful, but it's good stuff. Quite the group of experts assembled here on the phone today. This Tiger Team is going to be ably led by Dr. Helen Burstin. Helen, we really appreciate your efforts on behalf of this group and your leadership of it, so I will turn it over to you.

Helen Burstin – NQF – Senior VP, Performance Measures

Thank you, Jon. We have a kind of broad task without a whole lot of parameters around it, so I'm hoping what we can do today is specifically think about what are the methodologic issues that are likely to come up in the course of the Policy Committee selecting new measure concepts, thinking about specifically how for some of these we might even propose some solutions or probably more likely, if there are some of these measures that are coming forward, is there some additional work or study that needs to go on to allow us to really begin working with these sort of newer measure concepts going forward. The three targeted methods issues that were presented to us by Jon and Tom and Cray was the issue, first of all, of longitudinal measures over time; delta measures, which is really the change of a specific intermediate outcome, for example, like blood pressure, functional status, etc. over time; and lastly, the specific issue of adverse event reporting, where we begin looking at how we might be able to automatically capture some of these data within clinical workflow.

So that's a pretty tall order and I thought it might be useful—and if people don't think this is useful we could certainly try something different—that before we launch into each of those specific measure issues separately I often find it easier to deal with methodologic concerns when they're more based in something I can wrap my arms around, like measurement concept. Maybe it's because I'm just a measures geek, like many of you are, but I just pulled together a quick Tiger Team summary that Cray was good enough to send around to all of you earlier, where I took the nice slide set they had sent along to us, which listed out the leading measure sub-domains and measure concepts that each of the other content related Tiger Teams have been working on and thought it might be useful to just quickly walk through this to really make those three issues perhaps sing a bit more for us, so we can really get into the meat of which are each of those measures' issues.

I also specifically left the last column of methodologic issues not filled in and the hope is that even if we can't get through all of this today I would request that as you, over the next day or so, have the opportunity to really look through this in a bit more detail, if you could actually fill in some of these method's issues that you recognize will emerge from some of these key sub-domains and measure concepts and send them along to Cray that would be incredibly useful since I think we have all of about three days to present something on Thursday and would like to at least get your best thinking with how we're going to proceed.

Jon, ONC folks, does that sound okay?

Tom Sang – ONC

Yes. That sounds great. Helen, just for process matter, can we just do a roll call?

Helen Burstin – NQF – Senior VP, Performance Measures

I'm sorry. I thought that's what Jon just did. Okay. Sure. Let me see if I have the full list. Jon, do you have the full list of folks to do?

Jon White – AHRQ/HHS – Director IT

Let me dig it up.

Helen Burstin – NQF – Senior VP, Performance Measures

I have the list of who was on the last one, but I'm not sure that's the same thing.

Jon White – AHRQ/HHS – Director IT

I can do the list from here. I know we had Joanne Cuny from the AMA-PCPI. Do we also have Greg Wozniak on the phone? Okay. Joanne in place of Greg. That works. Mark Weiner I heard. Are you still there?

Mark Weiner

Yes.

Jon White – AHRQ/HHS – Director IT

Dan Malone from Arizona?

Dan Malone – University of Arizona – Prof. of Pharm. & Public Health

Yes. I'm here.

Jon White – AHRQ/HHS – Director IT

John Moquin from AHRQ?

John Moquin – AHRQ – IT Architecture Manager, Social & Scientific Systems

I'm here.

Jon White – AHRQ/HHS – Director IT

Ross Lazarus?

Ross Lazarus – Harvard Medical – Director of Bioinformatics, Channing Lab

Yes, I'm here. I'm going to have to leave shortly and then I'll be able to come back after about half an hour.

Jon White – AHRQ/HHS – Director IT

Okay. We'll take you for whatever we can get you. Bob Dolin?

Bob Dolin – HL7 – Chair Elect

Here.

Jon White – AHRQ/HHS – Director IT

Phil Renner from Kaiser?

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

I'm here.

Jon White – AHRQ/HHS – Director IT

Dr. Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT?

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

Here.

Jon White – AHRQ/HHS – Director IT

Karen Pace from the Quality Forum? Probably not, because we have two others—

Karen Pace – National Quality Forum – Senior Program Director

... yes.

Jon White – AHRQ/HHS – Director IT

Okay. Very good. Nita Roka?

Nita Roka

Yes, I'm here.

Jon White – AHRQ/HHS – Director IT

Yes. Abel Kho?

Abel Kho – Northwestern – Assistant Professor, General Internal Medicine

Yes. Here.

Jon White – AHRQ/HHS – Director IT

David Baker?

David Baker – Northwestern – Chief, General Internal Medicine Division

Here.

Jon White – AHRQ/HHS – Director IT

Tom Sang?

Tom Sang – ONC

Here.

Jon White – AHRQ/HHS – Director IT

Cray Noltert?

Cray Noltert – ONC

Yes.

Jon White – AHRQ/HHS – Director IT

All right. I think that's everybody.

Helen Burstin – NQF – Senior VP, Performance Measures

Thank you for doing that, Jon. What a great group, lots of good attendance today. So just in terms of an approach, does everybody have that Tiger Team summary worksheet that was sent around earlier today? Does anybody need it to be sent to them? Let me know. We can just quickly shoot it over to you.

Joanne Cuny – AMA-PCPI – Director

Karen, I might need it. If it came today it probably went to Karen Kmetik.

Jon White – AHRQ/HHS – Director IT

Actually, it would have gone to Greg Wozniak, but, Joanne, let me see if I can figure out how to get it to you.

Abel Kho – Northwestern – Assistant Professor, General Internal Medicine

Likewise for me as well, use my NMFF account.

Jon White – AHRQ/HHS – Director IT

Okay. You got it, Abel.

Helen Burstin – NQF – Senior VP, Performance Measures

Thank you, Jon, for taking care of that. Again, these are very high level concepts if you just want to pull up that document, but I thought it might be helpful to just run through the sub-domains and measure concepts. As we do so, please indicate if you agree that sort of fits into one of the targeted methods issues or does it bring up other methodologic issues in terms of how you would actually move these measures forward and have them work for meaningful use going forward.

Beginning with care coordination: Effective care plans, specifically the idea of getting clinical summaries and self-management plans; care transitions around transition elements; the patient experience of the transition potentially; appropriate and timely follow-up; medication reconciliation; follow-up labs and diagnostic results; patient follow-up after transitions; and this last categorization that actually I'm on that team as well and we're still sort of working through, is diagnostic management, the provider follow-up on the key issues that needed to be addressed, as well as communication with the patient.

I'm going to stop there. Anybody feel free to chime in if there are specific methods issues that those kinds of measure concepts, recognizing certainly you don't have the measure specifications in front of you, but it seemed like the concepts would potentially be very useful to illuminate the methods issue.

Bob Dolin – HL7 – Chair Elect

I don't know. I mean I sent to Jon and Tom over the weekend some what I thought were maybe over arching methodologic issues. I don't know how they, off the top of my head, fit into these buckets and I don't know what's the best way to try to raise them.

Helen Burstin – NQF – Senior VP, Performance Measures

Well, one thought might be—and actually, Jon just forwarded them to me as well—maybe before we dive too deep into the individual Tiger Teams, do you want to just briefly go over these, Bob? I think there is some really good information in here. Let's keep it sort of high level before we dive.

Bob Dolin – HL7 – Chair Elect

Okay. Yes. That would be great. There were six of them. Do you want me just to walk through them?

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. Briefly. That would be great. Actually, Jon or Cray, can we just get it to the whole team?

Jon White – AHRQ/HHS – Director IT

Yes. I will be happy to send it around. I also want to mention, Joanne, I actually don't have your e-mail. I'm sorry to everybody, but could you share that with me?

Joanne Cuny – AMA-PCPI – Director

Sure. It's my name, joanne.cuny@ama-assn.org.

Jon White – AHRQ/HHS – Director IT

Very good. On its way shortly. Thank you.

Helen Burstin – NQF – Senior VP, Performance Measures

Okay, Bob, why don't you go ahead?

Bob Dolin – HL7 – Chair Elect

So the first one was inspired by the Gretzky Report and Skate to Where the Puck Will Be: Here I just wanted to throw out this phrase of incremental quality measurements, because it struck me that some of

the things that we're potentially trying to measure may or may not line up with the bar being set by the meaningful use criteria for a particular year. So the more that we can keep the quality measurement and the meaningful use criteria for stage one, stage two, stage three in line I think the more likely we're going to be able to skate to where the puck will be.

Just as a for instance, here we are talking about longitudinal measures, but in the meaningful use criteria for stage one on the ambulatory side we require that longitudinal data be able to be reported in the summary, but that's not the case for the inpatient setting. If we were to now say, "Well, here is a quality measure that requires longitudinal data in the inpatient setting that would then have a disconnect to the meaningful use criteria," so the more we can keep these things step-in-step and kind of raise the bar incrementally across the whole spectrum, I think the more effective we might be.

The second kind of over arching issue had to deal with incremental e-measure, logic complexity. I think Danny and I have probably talked a lot about this, but what we find in the quality measures that we've been converting into the e-measure formulism is that 80% to 90% of the material is pretty straightforward and most of the measures only use this straightforward stuff, but there is a handful of measures that require extra bells and whistles. In trying to formalize the bells and whistles what you do is you raise the bar of complexity across e-measures as a whole. Just like we think in terms of incremental interoperability, it might make sense to think in terms of incremental e-measures, logic complexities, so that for instance, in 2011 or 2012 if we say we're only going to do those measures that address constructs that are in the CCD and that only use ... operators and only use averages. In 2013 we're going to turn up the heat and have additional expressivity in the e-measures, so that's one way that we might also think about how to go forward with some of these. That was the second one.

The third one was to avoid the need for retooling through proper initial tooling. What we find here is that—and I think this has probably been everyone's experience—when you have one group that's built a clinical practice guideline or that's built a quality measure or built whatever and they hand it over to another group, toss it over the fence, if you will, and say, "Here, go ahead and formally encode this," we inevitably wind up discovering areas of ambiguity and we have to go back to the original source and say, "What did you really mean here? What did you really mean there?"

I think a good recommendation coming out of the Policy Committee would be when you're sitting down to write a measure here are the kinds of people you want to have sitting at the table. This might include SMEs, the standards expert, the terminology expert, etc. to avoid the need to do subsequent retooling.

The fourth one was something we talked a little bit about last week, which is parsimony of data elements, not across settings, as described in the Gretzky Report, but parsimony of data elements across reporting requirements, because all of this stuff ultimately converges on the clinician working with their electronic medical records. So if we can say that the data elements we need for tuberculosis or the data elements we need for congestive heart failure are the same, whether it's for decision support or the same for a quality measure or the same for public health reporting, we can start to converge on data elements, the more likely it's going to be that we're going to be able to capture them at the point of care. That's obviously a place where the Policy Committee can take a strong stance and help direct the different federal agencies.

The fifth and sixth one are a little bit more obscure. The fifth one has to deal with shared object identifiers and I think the master patient ID one is a pretty common scenario where if you're trying to now look at these longitudinal measures or these delta measures and in your scenarios you have patients being seen by multiple providers that may be affiliated with different health plans I think the shared patient ID is the common issue, but a more complex issue is if you're trying to track certain conditions and you don't necessarily know that the condition provider A is referring to is the same condition that provider B is referring to because they're not using a common master file. In the current state of affairs with standards, where we have CCD and we're talking about here's the diagnosis, we share the code so we know that we're talking about the same SNOMED code, but we don't necessarily have a good strategy for sharing

and reconciling identifiers and so these shared object identifiers are going to become quite relevant if, in fact, we're going to be computing quality measures across care settings or across different providers in different care settings.

Finally, there was this one back when we developed the HL7 CDA Quality Reporting Document Architecture Implementation Guide we had discussed this role of a quality organization. So I thought it might help to point out that if, in fact, there is this notion from the Policy Committee about the role of a quality organization where that quality organization would be receiving data from the health plan or from Kaiser or whatever and from there cranking out the numerator and denominator data, we do need some type of recommendation around QRDA. Our work ... with QRDA would be that it would be able to convey sufficient data that would allow these quality organizations to be able to compute the numerator and denominator data, so we might need some kind of position there. That's it.

Helen Burstin – NQF – Senior VP, Performance Measures

Great, Bob. That's really helpful. A lot of those, I think, really resonate, at least for me, certainly as we think about moving forward. Any comments from any of the rest of the committee on what Bob set forward? Jon, thank you for sharing with everybody on the call.

Ross Lazarus – Harvard Medical – Director of Bioinformatics, Channing Lab

I'd just like to say Amen to all of that. I think those are all really big issues. I mean the last couple are, I think, the really difficult ones. Because of my public health perspective, obviously, I'm interested in assembling things like notifiable disease and to do that you have to get data from lots of places to do it reliably, because you need all of the tuberculosis tests that were conducted for a particular patient to find out whether or not they actually ever had a positive ... sputum. So I'd just like to say I think you did a great job and all of those things I think are really important, but I'm particularly obsessed about the problem of tracking all of the medical information that's available across multiple different providers, all of whom have a different ID for one individual patient and either a federated approach or a universal identifier, which, I guess in this country, is not going to be acceptable, but certainly the work that's going on in data federation is really, I think, directly relevant to what we're trying to do.

Helen Burstin – NQF – Senior VP, Performance Measures

Other thoughts?

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

I have sort of another conceptual issue to sort of pile on a little bit. I think one of the things that I'm struggling with is sort of a conceptual inconsistency, because I think that we're trying to build measures about a patient across the continuum and across settings of care within a structure or program that's designed to produce measures about a provider's interaction with that patient. So we're going to have to think about it in terms of the methods, whether it's the policy construct or whether it's how do we feed the continuum of care data back to each individual clinician so that they can then report it out of their EHR, but also so that they have that information so that they know. I think what we're really looking at is some sort of a two-way exchange to create kind of virtually integrated and interoperable records and so we may want to stage the measures in terms of we might have a structural; and I realize I'm talking more about what the measures might be rather than the methods issues, but I think that this might help us parse the different methods issues.

The first wave might be a structural measure, as in are you participating in an exchange, which we kind of already have. The second, we might have process measures about proportionate patients, whose data is uploaded or downloaded from the exchange and then you can use that to then get to the percentage of patients meeting goals. Each of these sort of has a different flavor of kind of methods challenges.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. Those are excellent points, Phil. Just a question for our ONC/AHRQ friends: Are we supposed to be dealing with the more structural side of actual use of HIT or are we really trying to still stick to the e-quality measures concept for this group's methods work?

Jon White – AHRQ/HHS – Director IT

I'll take it. My opinion would be that any issues that this wise group can foresee with accurately measuring quality through input from health IT systems, be it structural or otherwise, ought to be teed up on the table. Tom, if you want to focus them further than that let me know.

Tom Sang – ONC

I second your suggestion, Jon. I think the primary purpose is related to the quality measure work, but I think if you have secondary suggestions to the Policy Committee I'm certain they'll welcome any further comments that you may have.

Helen Burstin – NQF – Senior VP, Performance Measures

Great. Okay. Very helpful. Thank you. Other members of the group on sort of the general issues that Bob raised or Phil or just general issues of your own you'd want to tee up?

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

Yes. Just to follow along with Phil's previous comments, I think that one of the issues that we're sort of running into with e-measures are issues around attribution. A part of meaningful use is to sort of meaningfully use the EHR and in very many settings you'll have multiple providers carrying forward the same patient. In claim's based measures that was easy because attribution could be identified for who billed for what, but often times in an EHR system if you have shared care it's not always easy to track who entered what into the problem list, who is managing what on the problem list. So when you're looking at measures around a healthcare system or around an entire practice it's no a problem, but when you want it to get down to the individual provider and handle attribution that way that really requires a different style of measure or at least different style of information to bring along with every piece of data, who entered what, when, how, where.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

I'd actually like to respond to Danny's comment, because I was thinking about attribution as well. I think there are sort of two flavors of attribution. One is this is my patient and the other is I'm responsible for the outcome. I think that we've got some methods, some technical methods for assignment. They're in my EHR or I had X number of encounters with various sorts of tradeoffs and sensitivity and specificity.

I think the second, the one that Danny brings up, is as much a policy question as it is a technical or methods question or maybe the first question is a policy question of do we care who gets credit or blame or are we saying is it important that the patient got the care or not and everybody shares in, again, credit or blame? Then once we make that decision or somebody, not us, makes that decision then we may need to add the complexity that Danny raises, but my preference is to not get into that amount of detail in parsing it out, but saying that everybody has a share in it. But I think Danny is right; that if we go down that road that's going to be a very complex question.

Helen Burstin – NQF – Senior VP, Performance Measures

Right and we talked about this in the context of, for example, beginning to move for delta measures. If someone's A1c goes from 10 to 7 how does that get attributed? But I think the point you're raising, Phil, is actually not even just for the delta measure that this methodologic issue becomes apparent, but actually for all measures, even point in time estimates, without the ability to more clearly tie people to specific providers. It becomes a much more patient centric measure and perhaps how we use that becomes a little bit more complicated.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

And if you only give credit for doing something you can drive towards over use in terms of extra testing, extra; this is a silly example, but multiple flu shots.

Helen Burstin – NQF – Senior VP, Performance Measures

Ouch! Yes. Okay. Other thoughts on the same theme of attribution? No. Okay. Any other general sort of methods issues people want to raise that became apparent just as you were looking through the minutes we sent from last time or the comments from Bob before we sort of try to take a bit of a deeper dive into perhaps some of the Tiger Team domains and concepts?

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

I'll pile on with one more if you're not tired of listening to me yet. Danny or Bob mentioned a problem list. I don't know that we know enough yet about the sensitivity and specificity and positive and negative predictive values of the problem list, especially in the evolving world right now. We're making a lot of assumption, especially in the e-measure and the QDS, pointing at the problem list as saying whether or not somebody has a condition, but I don't know; this isn't an area of the literature I know intimately; but I don't know if we know if it's not on the problem list do they not have it. Does the patient not have it? We're really going to need to think hard about dueling problem lists. If they're just ..., if one list says the patient has diabetes, but another clinician doesn't, how do we resolve that in saying is this patient in a diabetes measure?

Helen Burstin – NQF – Senior VP, Performance Measures

Right. Actually, David Baker, if you're still with us, I wonder if you want to share your experience with the work you guys did at Northwestern looking at, for example, what was documented around Coumadin use, as I recall.

David Baker – Northwestern – Chief, General Internal Medicine Division

For Coumadin lists I mean the medication, what's under reported are the exceptions. The problem list is interesting and I've looked at this and Abel Kho has looked at this as well from a very different perspective and we're familiar with some different partners here. I'll just say, in answer to Phil's comment, it's hugely variable. The accuracy for hypertension and diabetes on our problem list is very good. We know of one of our partners here, who it's very bad, so bad that they actually set up pop-up alerts to remind physicians to add it to the problem list using an algorithm based on the medications that the patient was on and some other things. But for myocardial infarction, for patients that we know had an MI, it's only on the problem list about half the time. So we have to drive our quality measures off old encounter diagnosis codes, because people tend to put it in when they see somebody after an MI.

I'll say that part of the reason for the variability, some of it's just experience with the systems. Some of it is just the nature of the way we care for people, but some of the systems that we've looked at, the problem lists are clunky. Putting things on the problem list are clunky. You can't append well and put notes into the problem list, so it's always going to be hard to get people to use the problem lists in the systems.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

I wonder there too, Dave, I mean used to struggle with this; if a patient's had a heart attack and they also have atherosclerotic coronary artery disease and they have angina, struggling to know do all of them go onto the problem list, do I abstract it to this one? Then based on my decision, if the measure comes along and if they happen to glom onto the wrong set of codes will they miss the way that I happen to reflect it? I mean did you find any of that?

David Baker – Northwestern – Chief, General Internal Medicine Division

I think you're exactly right. I mean clinicians tend to document the thing; they use the problem list as a tool if they're using it and if somebody has active angina I'm going to have that on the problem list. If

somebody had angina and they were revascularized and they have not had angina then I'm probably going to put coronary artery disease. So this just speaks to the need to be able to, in an ideal world, try and have systems. The rule is if somebody has had a myocardial infarction in the past you put it on the past medical history and you use the past medical history and the problem list codes to drive your clinical decision, support and your quality measures.

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

David Baker – Northwestern – Chief, General Internal Medicine Division

But the other thing I will add, I mean there is some data suggesting that if somebody has diabetes and it's not on their problem list or they have heart failure and it's not on their problem list, they actually are getting worse quality of care and there may be a causal relationship. How we solve that problem to get people to use the problem list is a difficult one. I think there's going to need to be a lot more work to understand the ideal way of defining the denominator population, as well as the ideal way of setting up systems to remind us when their problem list or other things are out of date.

Helen Burstin – NQF – Senior VP, Performance Measures

That's a really good point, David. Actually, if you look at the list of some of the sub-domains and measure concepts that were brought forward, for example, the Efficiency Team specifically talks about a longitudinal dashboard for leading conditions. Clearly, this issue of what do you rely on to get a sense of who fits into which of the leading conditions is going to be a pretty important methods issue to work through. Any thoughts, perhaps Tom or Jon or Cray, let me know if this is off base, but any thoughts of how we would actually begin to do some of that background methods work to help ONC and others understand how to make that doable?

Cray Noltert – ONC

That's not off base, by the way. You're heading towards good suggestions for how to address the issues, so carry on, please.

Helen Burstin – NQF – Senior VP, Performance Measures

How do we actually do this? How do we actually try to begin to sort of get some of the information we need to better understand these methods issues and perhaps build in different kind of guidance to, as Bob was talking about, tooling measures? It does change the way we begin to conceive of the way we develop measures as well for people like Joanne and others, who are on the line, who are measure developers.

Mark Weiner

Perhaps what this is suggesting is the need to take on another one of the ONC meaningful use criteria, which is interoperability, and really going outside the single system for making the call about whether or not a diagnosis exists. I think we can get into some debate about what should be the ultimate source of truth for certain diagnoses, where there might be some debate if it's there or not, but frankly, if there's an inpatient admission corresponding to a patient in my ambulatory system where it says there was a heart attack I would trust that over and above the presence or absence of a heart attack diagnosis in the problem list of the ambulatory system.

David Baker – Northwestern – Chief, General Internal Medicine Division

\ That's a great example and in an ideal world for some of these things, for something like myocardial infarction that probably has extremely high predictive value, you could even force some of these things on the problem list or at least give the physician a tool so that they're advised that they should add it in. MI is probably the best example, as I say, of something that's really poorly represented on the problem list but, as you say, very accurate on discharge diagnosis.

The problem is though that we found a lot of these people, they have the heart attack outside of our system, which speaks to the issue about the interoperability, but it's there on the physician's initial note, so some of this is setting it up so it actually meets physicians' needs. There's an easier way of documenting these things as people gain more experience. Some of these things may be helpful. We thought about for some of them that maybe Abel can speak to some of these algorithms that have been set up for diabetes, but it's really tough.

Abel Kho – Northwestern – Assistant Professor, General Internal Medicine

Yes. This brings up a whole bunch of issues around scope too. I mean I think people touched on it with the attribution question. I mean if we're going to be talking about going outside an individual EHR to sort of more of a broader health information exchange model it will be tricky because the meaningful use measurements are tied back to the individual provider. So I think we have to be sort of careful about maybe not initially, maybe we still need to sort of focus it back into the individual sort of level of provider and just do a better job about defining attribution potentially.

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

Bob, can you tell us about any standards work that has been done around recommendations for how a problem list should be used? The reason, Bob, I'm asking you that question is a lot of emphasis, and rightly so, is being placed on the problem list, yet I think if we look across the board for even people on this phone call how we all use our various problem list, depending on our location and on our EMR, we're going to get very different things out of it. It has factors such as usability, structure of data and even mostly local culture. So I wonder if there have already been recommendations on these are the ways to most meaningfully use your problem list.

Bob Dolin – HL7 – Chair Elect

There has been some. I'm trying to remember where it was where we debated what goes on the problem list versus what goes in past medical history. I'd have to dig it up. The other thing that we used for a precedent was, I think, Joint Commission or was it CMS recommendations for how to manage the problem list? But there is nothing in HL7 that I know of, Danny.

Helen Burstin – NQF – Senior VP, Performance Measures

One thought also is some of the work that Danny and others and Bob is doing around the quality data set and e-measure format has been this whole effort, for example, to categorize a problem as being active versus inactive. But the question is I think what we're hearing here is that it sounds good, but actually how does that really happen in the context of EHRs to be able to make that, to have greater assurance that the diagnosis you're using is the basis for this longitudinal dashboard is appropriate.

The second this is AHRQ has done a conference, just a couple of months ago; Jon was at it as well; looking at how do we begin thinking about how we bring together claims data plus EHRs. I'd be curious to hear some of the methods issues of sort of that more blended model.

Jon White – AHRQ/HHS – Director IT

Helen, that's a great point. My binder for that is in the next office over being reviewed by my staff that's on your HITECH Act, so I will absolutely bring that and the subsequent discussion back into what we bring forth Thursday if that's okay.

Helen Burstin – NQF – Senior VP, Performance Measures

Sure. Sure.

Joanne Cunny – AMA-PCPI – Director

Helen, I'm kind of a new voice to this group, but I'm very interested in what you were just discussing about the difference between medical history and the problem list, because you said for the problem list that's an active condition or is it an active condition. Then I reflect back on what Dave said; I think it was

Dave Baker that said that there's a patient that's had a previous infarct, but there's no reference to cardiovascular disease on the problem list.

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

Joanne Cuny – AMA-PCPI – Director

So it makes me think what's active. I mean I'm wondering; the physician is saying what's pertinent to what I'm seeing this patient for today and they don't think of a myocardial infarction as being active because it's in the past history, but still, that patient has coronary artery disease, so I would say that that is a big problem in trying to give good direction in how to use the problem list because coronary artery disease does play a part in what I'm treating the patient for today, probably regardless of what the patient is here for today. But it's very difficult to get all of those chronic conditions, if you will, onto the problem list when, in fact, they do play a part.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

It even makes me think of a topic near and dear maybe to me and Danny's heart, which is this notion of standardizing the EHR context. It would be great if you could say here's a piece of information I want and here's the place in the chart I expect to get it. I think most electronic medical records have the notion of a MAR and a problem list, etc. The more that we could drive towards having a standardized set of contexts, the less work I think it would be for all of these people to have to find the data, grab the data. Specifically here, if we say that problem list is one context and the past medical history is another context the more clarity we could put around the rules for which gets populated where. I mean I remember at Kaiser this was an ongoing challenge and there was never a lot of consistency. If I have something on the problem list and I change its status from active to inactive do I therefore then go ahead and put it on the past medical history? Well, typically no. You would just leave it on the problem list as inactive.

David Baker – Northwestern – Chief, General Internal Medicine Division

I agree with what you just called for, Bob. I agree completely. I think the next step though is if physicians and other providers understand that the problem list is a quality improvement tool and they understand that their clinical decision support, which should help them in terms of efficiency as well as quality, is all driven off the problem list, then I think people will start using the problem list more. So it's that issue about standardization, but then also making it more useful. That's where we've been trying to drive the physicians, to really see this. It's been a long process, but we're at the point now that some of our quality measures and some of these other things, these essential medication lists, that we sent to physicians, we stopped using old encounter diagnoses, because now the problem lists are better and these old encounter diagnoses, now we're dealing more with false positives rather than the false negatives that aren't on the problem list.

The key thing is docs need to have some advantage by keeping the problem list up to date and there needs to be a simple way to be able to deal with resolved problems and deleted problems. That's the way Epic is set up. Somebody puts a diagnosis onto the problem list and then it turns out it was actually incorrect, which all of the clinicians know happens, so they can delete it. But then we also have things like patients with diabetes, who have lost weight or some of them have had bariatric surgery and they literally no longer have diabetes and docs are getting reminders and it's part of their performance measures for these diabetes measures. So they can resolve a problem and it doesn't appear on their active problem list. We can decide for each measure how we want to handle the clinical decision support for resolved problems, deleted problems and active problems.

Dan Malone – University of Arizona – Prof. of Pharm. & Public Health

I just wanted to chime in on the question about should you populate problem lists with electronic claims data or encounter data. The answer, I guess, is maybe. One of the issues that you run into with populating that data is that you get sometimes a degree of specificity that reflects inaccurate precision

because the diagnoses codes, somebody could pick something that doesn't quite fit, but yet is too precise for the general condition that we're trying to capture there.

The other issue is just issues with miscoding and up coding—well, maybe not so much up coding, but miscoding being more of a problem there. Then, as we just heard, the issue about how long do you keep that going on the problem list, so it's really kind of difficult to merge those two, although one could argue for doing it and not doing it.

Helen Burstin – NQF – Senior VP, Performance Measures

Those are great points, Danny. The other question I'd bring up is as you make the transition away from ICD-9 to ICD-10 or SNOMED we may get a much greater degree of specificity in the ability to track these data. I mean I'm always frustrated when I'm at clinic and I have to pick a diagnosis code and all I can find is pain in foot or something where I could be much more specific if the codes were there. Any thoughts about the shifts towards the different coding systems would be welcome as well? Because I think that's going to be a pretty significant challenge as we move towards ICD-10 over the next couple of years, as well as SNOMED.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

Yes, Helen. One thing that did occur to me, as you and I think it was Dan were saying, especially as we've done a lot of work over the last ten years on how do we impute diagnoses or conditions from claims data into sort of an ICD-9 terminology. I don't know that we know how to do that into ICD-10 or SNOMED. As was mentioned, there's a degree of precision to those terminologies that we could be sort of building false precision into that without really knowing how we're matching those things up.

Helen Burstin – NQF – Senior VP, Performance Measures

Right. Other thoughts on that area? Otherwise, if not, I'd like to kind of keep us moving on this longitudinal issue, but a slightly different vent. We've mainly talked about the issue of how do we code to ensure we know what the condition is, for example, or the past medical history, which are critical issues, but if you actually look through some of the measure concepts that have been put forward by some of the other Tiger Teams there are issues like follow-up on lab results and diagnostic results, follow-up after transition, not ordering lab repeat lab results when they weren't necessary. So those are slightly different views on longitudinality, but the ability to look across systems and providers. Any thoughts of the key methods issues there, which seem endless from where I sit, but we'd really welcome your thoughts there?

Mark Weiner

One thing we wrestle with here is when we see a first diagnosis of diabetes. It's often hard to figure out if it's new just in our system or it's new to the patient. We come up with various algorithms for trying to figure that out, so like if a patient has been in the system for several years and there was no code for diabetes and suddenly there is a code for diabetes, we're more apt to say that's a new diagnosis; whereas, if a person is brand new to the system and on their first visit they get a code for diabetes then most of the time when you actually read the chart the guy has had a 20-year history of diabetes and we're just seeing him for the first time.

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

Mark Weiner

But that has implications for your expectation of how with the program the patient is from the get go in terms of their having a solid understanding of the medications they're supposed to take and getting all of their eye care and foot care and all of that other stuff.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. Okay.

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

I think that also dovetails well with our previous discussion for managing of the problem list. Most EMRs have a way to indicate the date of onset for a given disease process. I think that most providers just sort of skip over that and it defaults to whenever it was entered. If we could have some guidance and resolution on how to appropriately use the problem list I think that would help out with the sort of new versus recurring episodes.

Mark Weiner

Right. My excuse for that, because I do that all of the time, is when it asks for start date it's always a date and I don't know when it started, but I can say five-ish years ago and having a different way of having more of a fuzzy start time I think would be more appealing to people and allow people to fill it out more honestly rather than saying January 1, 2005.

Helen Burstin – NQF – Senior VP, Performance Measures

Okay. Great.

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

So is any one of the things that we're suggesting already in the sort of stage one meaningful use they talk about maintaining this up to date list of certain active diagnoses. Is it that we come up with additional guidance about how to really maintain or what sort of details need to go into the problem list?

Mark Weiner

Well, the C-32 spec is interesting, because it references module calls, active problems. The other thing I don't recall off the top of my head is is there sufficient guidance right now in C-32 regarding the date at which the problem was entered into the record versus the date at which the problem began with the patient? So this is kind of getting back to the Gretzky theme of skate to where the puck is going to be. If this is critical for measures then that would be something that I think in the next iteration of C-32, maybe as part of stage two of meaningful use, we also want to make sure that the module is fleshed out in that area.

Helen Burstin – NQF – Senior VP, Performance Measures

That's very helpful. So, beyond the diagnosis stuff, which obviously is critically important here, how about thinking more about sort of information and I guess it's more about interoperability, but how would you begin to look at measures, for example, that look longitudinally? There is an example under the population health/healthy lifestyle behaviors of looking at the longitudinal effect of smoking or looking longitudinally at whether patients got follow-up after a transition. Methods, issues that arise, those kinds of longitudinal issues, as opposed to the diagnosis specific focus?

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

Yes. The main issue that I see around longitudinal evaluation, this may be ... but some consider data our data and for longitudinal analysis it just requires data across settings and I think maybe 80% of the longitudinal and methodological challenges can be resolved with interoperability of information across systems, provided that the necessary sort of modified information travels with those data.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

Mark Weiner

I agree with that statement. I think some of the measure concepts here lend themselves to good data collection, especially if you can get data from across hospital systems, but a couple of them I see all of the time and I always cringe; one of them is the self-management plan and then the other is medication reconciliation. The reason is I think these kinds of concepts fall into the spectrum of what I call quality by checkbox, where it's there or it isn't there, but what really matters is how good the self management plan was conveyed to the patient. Really, you cannot get that from an electronic record at all. The best you

can do; we gave the patient a report card that showed all of their quantitative information and gave them a piece of paper with how they should vary their insulin regimen depending on what their sugar is, but I see a whole range of patient styles that make the gift of that information not equally absorbed.

The same with the medication reconciliation. I mean I do all of this work. I add medicines. I remove medicines. I clean up the medicine list, but I don't get a credit for it unless I click the reviewed box on the electronic record.

Jon White – AHRQ/HHS – Director IT

I think one big issue I see around longitudinal measures is I think it still goes back to an attribution problem. I mean if there are, let's say, three systems in town and a certain portion of the patients go back and forth between all different hospital systems, who is going to get credit for those guys at the end of the day and who is going to go through the trouble of trying to pull data across all of these different systems? Even if they are interoperable does an instance of that data have to be at every single system or are we essentially driving towards more of sort of larger hubs that aggregate the data on behalf of those institutions? So I think that can be kind of a tricky space.

Helen Burstin – NQF – Senior VP, Performance Measures

... more of a community HIE model.

Jon White – AHRQ/HHS – Director IT

Right and that becomes the aggregator of the data and actually does the quality measures and does it on behalf of the systems, but at the end of the day you still have to have some pretty clear definitions around essentially whose patient is whose.

M

Yes.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

When I was at NCQA I mean we really struggled with the medication reconciliation because you're basically trying to measure a cognitive process. I think that some of the self-management and communication with patients, that's tough. Maybe there are a couple of other ways we can get at it. One is requiring this exchange, but the degree of concordance between med lists across providers, rather than having sort of ... med lists.

The other thing for some of this is I wonder if we can get at it through patient experience, which, when we get there I've got a whole long list of other issues to talk about, but asking the patient did you understand your medication list, did you understand what you needed to do when you went home from the hospital rather than trying to get inside the physician's head.

Helen Burstin – NQF – Senior VP, Performance Measures

Other thoughts on that line?

Jon White – AHRQ/HHS – Director IT

We've got pretty good data that shows people don't know what they don't know and they basically give the healthcare team credit if they were nice and they spent time with them. So there is no perfect solution. I actually agree that the patient experience may be one of the most practical things, as well as the most important. I agree with the comments that were made about the checkbox for the medication reconciliation; that people check the box and they never went through it.

You think about all of the patients, for example, who are discharged from the hospital and their first follow-up is with a sub-specialist. The sub-specialist can't even really do any medication reconciliation for the nine other medicines that are the non-cardiac medicines, because they didn't prescribe them. Sometimes the hospital discharge summary is not correct, so again, I wish I had more solutions, but I agree the patient experience is an important part of the equation. It would be great for the medication reconciliation if there were good tools and some of the electronic health records are developing them and then it's going through those tools and leaving a mark rather than actually just checking a box that said that you did something when you really didn't.

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

M

Helen, it sounds like we've identified almost two definitions for "longitudinal." It sounds like one definition is the use of sort of traditional slice in time data, such as a single blood pressure reading that is measured over a period of time versus the sort of second method, which is using slice in time evaluations to assess a longitudinal outcome, such as the patient experience or even things like medication reconciliation. That's still done at a slice in time, but that gives you an idea of some longitudinal information.

Helen Burstin – NQF – Senior VP, Performance Measures

I like that.

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

The first bit can be solved with interoperability. You've just got a whole bunch of blood pressures across settings.

The second part, looking at the longitudinal individual data points that assess longitudinality of the whole process, those you don't really need interoperability. Those you just need effective assessment instruments.

M

I agree with what you said. The one exception, I think, is these patients, who are discharged on a medication and then it's not continued at their follow-up, so that's really sort of a longitudinal measure where I think you do need to have the data for

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

Sometimes I think that that requirement is a little over stated and within our own hospital and our own ambulatory practices the hospital doesn't necessarily have a grasp on certain formulary issues, so they'll often get discharged on a medicine that the patient has no prayer of being able to pay for. I do agree with the importance of the timely follow-up with the ambulatory care doctor, because that's a point in time where adjustments can be made, but I don't like this point-by-point reconciliation where what they were discharged on is the gauge for truth.

Helen Burstin – NQF – Senior VP, Performance Measures

Why don't we stick on that specific topic of, for example, the timely follow-up, the timely follow-up with the ambulatory care provider after discharge or timely follow-up of their abnormal labs where they followed up on things like that? Thoughts about specific methodologic issues there? It sounds like, obviously, a huge issue is this issue of interoperability, as Danny and others just laid it out, but beyond that, other issues, more methods issues around measurement in those areas that perhaps go beyond interoperability or perhaps going back to Bob's point about incremental measurement here. Is there sort of an opportunity to think about some of those effective assessment tools Danny talked about versus having a set of measures that truly require interoperability that perhaps may still be too much of a reach in the coming year?

M

One of the ones I think is really important for us to be thinking about is that time to first follow-up visit. That may be one where the claims data, as was discussed previously, becomes really important. The unit of attribution for that really, which is ... to question, is this the primary care physician's responsibility or the hospital or both? If you have the claims data and you can identify who the primary care physician is or the sub-specialist, then you should be able to attribute it to both and track that; whereas, if you just have one system it might be difficult.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. Other thoughts?

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

Yes. At the risk of joining Danny and the ... box, I'm not sure that some of this can actually be done well out of the EHR from a pure measurement perspective, things like readmissions, these timely follow-ups almost are going to be more efficient out of claims.

The question though comes back to how do we support it, because the claims we can then go after and sort of ding the clinicians, but how do we use these tools to build systems to sort of push follow-up and kind of build a forcing function, which may be some of the more structural or functional measures within stage two or stage three rather than sort of outright clinical measures. Does that make sense?

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

M

Phil, what if, to play devil's advocate, in your EHR you did have access to information from other EHRs? Imagine an EHR integrated within to a health information exchange. You could make the assumption, which is a very, very far stretch, that you'd have access to HIE data. Do you think that longitudinal measurement could be done?

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

Yes. Okay. I think it could be and I think that that would be the way to support this. I think it's a big assumption and one of the things, one of the methods issues we haven't gotten to yet is around data completeness. We were sort of dancing around it a little bit with the problem list, but if that data isn't being pushed or if the hospital doesn't know who to send the discharge summary documentation to, I don't know from a measures perspective if that's an exception or if that's a problem, if that's sort of a negative and sort of who, in that case, who gets dinged. But yes, right, I think that the exchanges will help that.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. We probably can't solve that issue today. It's kind of a big one. Given the time, we've talked a fair amount, I think, about some of the key issues in longitudinality and by no means have we covered it. Perhaps at the end we'll come back to perhaps a home exercise you could help us with thinking more deeply about these.

The second big methods issue that was keyed up for us as part of the work we had done with the Gretzky Group and the e-quality measures report was this idea of not always just moving towards quality measures that have a threshold value, but also using the functionality of the electronic health record to move towards being able to look at the trajectory of care over time. For example, if we want to play one out, if we were able to look rather than pick your A1c thresholds de jure of 8, 9, 7, but instead could track A1cs over time for all patients how would we begin constructing measures differently and the methodologic issues involved and using the data in hand to begin thinking about looking at that delta measurement over time. Does anybody want to start the discussion there?

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

I'll throw out a few things and I'll push back on the question of not needing a threshold. I think we're talking about a different threshold—

Helen Burstin – NQF – Senior VP, Performance Measures

I agree. Yes.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

Right. So you're looking at either some sort of absolute or percentage change. The threshold might be you've got to move blood pressure 10mm—

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

It might be a threshold around percent of the gap to target closed. It might be the percent of patients that are getting any improvement, even if it's minimal. Then there are a couple of things, assumptions, under this. There are a lot of measures we talk about where we're saying there is more benefit to moving somebody who's way out, has really high blood pressure to just kind of high blood pressure, but the gradient may not be constant. I don't know the degree to which we've sort of mapped that gradient and sort of how much credit do you get for the 10mm that are really high versus the 10mm that bring you just above threshold.

Then we've got the J-shaped curves where if you bring blood pressure or A1c too low you're increasing risk again. Then there's what if the patient doesn't move monotonically in one direction, if they kind of bop up and down? What do you do there? Those are my questions. I don't know what the answers are.

Helen Burstin – NQF – Senior VP, Performance Measures

That's a great set of questions, Phil. Thank you for starting us off. Other thoughts?

Mark Weiner

Related to that is just symptoms that come up, so the patient's blood pressure is 150, but when you try to lower it more you can, but the patient is dizzy or the patient has too many hypoglycemic episodes with diabetes, so –

Helen Burstin – NQF – Senior VP, Performance Measures

Okay.

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

To add on to Phil's comments, all of the methodological calculations are technically possible. As Phil was suggesting, one of the questions is going to be completeness of data. This goes back to the issue of interoperability with longitudinal data, so how do you ensure? Let's just say that you are pulling records from a couple of different systems. How do you ensure that you are including all of the data points?

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

David Baker – Northwestern – Chief, General Internal Medicine Division

My biggest concern with some of these is will these change measures actually be valid indicators of quality of care for the reasons that have already been mentioned. Some people you do everything right and they can't tolerate their third or their fourth anti-hypertensive—

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

David Baker – Northwestern – Chief, General Internal Medicine Division

And you haven't done anything wrong. Actually, that's a big percentage of this group and we've been looking at some of these things. It may be, if you think about it, for a patient centered medical home model it may be to say somebody who has a blood pressure above 150 repeatedly, that that person should be referred to a nurse practitioner or pharmacist or some other system based solution.

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

David Baker – Northwestern – Chief, General Internal Medicine Division

I just really worry that this tracking A1c or tracking blood pressure over time, I worry about the validity of that as a quality indicator.

Helen Burstin – NQF – Senior VP, Performance Measures

That's an excellent point, David. I think part of where this also came from, and some of this happened as part of the database purchasing program CMS had done for some of the small, rural hospitals who, perhaps, couldn't achieve the threshold, but were making significant gains towards getting there. I think part of this is trying to reflect on the range of providers out there. If, for example, a patient who comes to my health center newly diagnosed with diabetes, comes in with an A1c of 14 or 15, which we see unfortunately, I can get them down to 9, which is probably phenomenal in terms of their impact on health status and yet, if the threshold measure is 8 it will look like I'm out of range. So I guess it's not an absolute you would go for one over the other, but does this give you the flexibility to allow for improvement along the trajectory?

David Baker – Northwestern – Chief, General Internal Medicine Division

That's one way of looking at it, but we've looked at patients, who repeatedly have an A1c above 8, so they're 8.5 and 8.6 and they're not heading downward and we've done chart reviews on small samples of those and the doctors are responding. They're having discussions with patients and patients are saying, "I want to try lifestyle modification. I want to try these other things," but we don't have any system level support really to helping them with the behavior change and most of the people won't be able to do that anyway. So our impression from that was we really need to develop better systems support to be able to help people and then hopefully, if they don't succeed in that, even with formal support, then maybe they'll accept insulin.

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

David Baker – Northwestern – Chief, General Internal Medicine Division

But again, if I was going to look at everybody in that group and say were they getting good quality care based on current resources I'd say yes, most of them are.

Ross Lazarus – Harvard Medical – Director of Bioinformatics, Channing Lab

Just come back and the A1c issue in change terms is really exciting and it's a very interesting conversation, but I mean how many people are actually having it measured is something that we don't yet know. I mean we're working on another project where we're looking through data from a particular provider and the number of people, who are having A1c measured is depressingly low. I'm just wondering if before we talk about change measures or longitudinal measures that needs to be staged, because if we don't know that it's being measured, if we don't have data that says something really fundamental, like this proportion of people over 40 had an A1c measured it's really difficult to know how far we can get to getting change measures, let alone longitudinal measures.

So I think this is all a really interesting discussion, but it kind of assumes that we've got the data to make that quality measurement. Once we've got the data then we can enjoy all of these really difficult

problems, but until we actually know how many people are having it measured I don't know that it makes a lot of sense to talk about change measures.

Helen Burstin – NQF – Senior VP, Performance Measures

I guess that's the difference of a population perspective versus at least knowing among your patients who is getting it measured, but it's a very valid point.

David Baker – Northwestern – Chief, General Internal Medicine Division

We actually, if somebody has not had it measured, then they fail the quality measure on a cross sectional basis. If somebody has one that is out of range and it's not dealt with within a certain period of time then you could say that that person failed the measure. I think that's important because it would encourage people to set up systems to identify people, who have fallen out of care and do outreach, because that's what we see in our system. I always say our biggest quality of care problem is the patient we're not seeing and we're just beginning to set up systems to identify those, who have fallen through the cracks and do outreach.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

Joanne Cuny – AMA-PCPI – Director

I have a question; this is Joanne; about those same patients. Do you, in your system, have a method where you can readily identify a patient that may be an exception? In other words, that refused insulin or is non-compliant. How do you capture that?

David Baker – Northwestern – Chief, General Internal Medicine Division

We have very simple tools that the doc use, particularly for medications and other things, but I'll tell you, it's interesting. For the one for diabetes, the doctors are very reluctant to use that because they want those reminders to keep coming up because they're still working with the patient. Right? They don't want to just get credit and check the box. They want the decision support tool to say, "Hey, this person is still out of control," and if they just say patient refused then they'll only do that as a last resort and we think that that's appropriate.

Joanne Cuny – AMA-PCPI – Director

So if they document that the patient refused then they come out of the decision support?

David Baker – Northwestern – Chief, General Internal Medicine Division

If a patient refuses, let's say they refuse pneumococcal vaccination and typically the docs will check off that the patient refused and the alert is turned off for one year and the doc will address it again. They do that all of the time. It's very common for medications and preventive services, but for these intermediate outcomes it's very few times. It's usually maybe after two or three years the doc will finally say I've been trying to get this person to take insulin for two years and I'm just going to turn this off. But it's a challenge. We really want people to only turn those things off when they don't want to be notified any more.

Joanne Cuny – AMA-PCPI – Director

Right. Yes.

David Baker – Northwestern – Chief, General Internal Medicine Division

So what we really are trying to do right now on a pilot basis and we've just started this in the last few months is identify those people, who are persistently out of control and either do outreach directly or ask the physician if they'll refer the patient in for disease management. You could imagine a system that says if the person is out of control have they been brought back into control or have they been referred to this outreach program. That's as much as you can do. It's intensive, one-on-one, motivational interviewing and if the person fails that you've done everything you can, so it's sort of a hybrid measure.

Joanne Cuny – AMA-PCPI – Director

Right, which makes it a very complex measure. That's actually what I'm getting at. That makes it very complex to use it as a measure.

David Baker – Northwestern – Chief, General Internal Medicine Division

It actually wouldn't be that complex really. We could do this very, very easily. It's just you have to know the code for the referral to the diabetes educator in that case.

Helen Burstin – NQF – Senior VP, Performance Measures

Other thoughts on the delta issue? Does it add much?

David Baker – Northwestern – Chief, General Internal Medicine Division

Could I just mention I think the delta thing that's the most exciting to me that you referred to is the change in health status. I'd love to have vision before and after cataract surgery and after hip replacement. The cataract surgery one, every patient going for surgery asks me who is best in our system and I always say the same thing, "I have no idea." So I think that's a really exciting area. It's still challenging, because it may be you're going to have to really push the surgeons or have them work with the primary care physicians to get that data—

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

David Baker – Northwestern – Chief, General Internal Medicine Division

But I think that's one of the areas we should really be thinking about.

Helen Burstin – NQF – Senior VP, Performance Measures

Other thoughts there?

M

One other thing under the methodological thing. I think it's great if you can get someone's A1c from 13 to 9 and I think a ... change measure would capture that in the first year or the first two years, but how to handle the issue four or five or six years later when their A1c was horrible and now it's better, but it's still not threshold. Are we still aiming for the A1c to become less than 7 in the person who already had their drop of a significant percentage and now they're sort of stuck in the mid-range?

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

David Baker – Northwestern – Chief, General Internal Medicine Division

So if the patient's A1c stays at 13 and they have four visits with their primary care physicians for diabetes as an encounter diagnosis over the next three months but no change is that bad care?

M

No.

David Baker – Northwestern – Chief, General Internal Medicine Division

I've got plenty of examples from my own practice where I've done everything that I possibly can and they're not making any improvement. I think, at least in our system, that's a big chunk of those patients that would appear to fail a measure of improvement in A1c. I don't think that they're getting bad care.

M

It kind of depends on what the purpose of doing the quality measurement is. If we're going to be basing pay-for-performance bonuses on it I do worry about that, because then you're penalized for taking care of this guy where it sounds like you're saying it's not on you that the A1c is remaining 13. But if we are

trying to define a population that might be in need of additional support services then I think that does matter. I agree; I wouldn't want to ding physicians for taking care of these patients, because then they'll just move to the rich suburbs where, on average, patients do better and ... take care I'll get off my soapbox now.

Helen Burstin – NQF – Senior VP, Performance Measures

There are plenty of us who could replace you.

M

I know ... phone call

M

Helen, you did actually ask something or you started to. You said, "Is it worth it?"

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

M

I think that an area of testing would be from a measurement and sort of information availability perspective does doing this give us any more information in terms of rankings or scorecards than the thing we have now. So we could be just talking about adding complexity without adding value.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

Bob Dolin – HL7 – Chair Elect

Helen, I think it was maybe in one of the e-mails that was sent out after last week's session, the summary of what some of the methodologic issues are for these delta measures are, first off, figuring out what to measure. Do we measure averages? Do we measure blah, blah? What's the position on how do we tie back this to the clinical evidence, to the clinical trials, to the evidence? It struck me a little bit odd that we had to ask ourselves what do we need to measure. Do we need some type of methodologic issue that ensures that what we're measuring is what we've identified based on evidence to have the biggest impact?

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. I don't think, at least I wasn't if I was the one who said that; I was intending it to be that we don't know what the mean should be, but more so, for example, who determines what the two points in time are, issues like that. Going back to the issue we started with, is it when you enter care? Is it the average of the year before to the next year? Just those kinds of issues I think are not clear if you consider moving to those kinds of measures. Again, I don't know of many places that are doing these yet either, those issues specifically highlighted by many of the ONC folks as well.

Tom, do you want to add anything here? I know this was a special issue for you and Farzad early on.

Tom Sang – ONC

No. It's just, Helen, when you ... the Gretzky ... this issue was really highlighted from the folks on the Gretzky Group—

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

Tom Sang – ONC

So we're actually taking your advice and the group's advice and the group of subject matter experts on quality measure developments that this is a robust data point that we should consider that could help in quality improvement.

The other thing I want to point out to the group is a lot of the issues that you're pointing out can be parsed out either to the Policy Committee or to the Standards and Certification Quality Sub-Workgroup where they would then talk about the standards involved. Then the other issue is that things that may not necessarily reach an impact in stage two can certainly be considered aspirational goals for stage three.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. Okay. Got it. Thank you. Certainly, the health status stuff was the part that was the most of interest to many of the folks longitudinally across conditions, which we talked about. My understanding from our last call is that those methods issues are being handled currently by the Patient and Family Engagement Sub-Group. Is that right?

Tom Sang – ONC

That's correct. Yes.

Helen Burstin – NQF – Senior VP, Performance Measures

Very good. So, any other thoughts on using measures that sort of rely on a delta before we move on to any other methods issues? Certainly a rich discussion. Lots of pluses and minuses, as we've heard.

David Baker – Northwestern – Chief, General Internal Medicine Division

I would just add—and I know this is way beyond this group's scope—but I would love to see for some of these procedures like cataract surgery or hip replacement that people just don't get paid until they have completed some health status measures. I think that's the only way you're going to get people to actually report this data.

Helen Burstin – NQF – Senior VP, Performance Measures

....

M

Yes. I would add back surgery to that list.

David Baker – Northwestern – Chief, General Internal Medicine Division

Agreed.

Helen Burstin – NQF – Senior VP, Performance Measures

... some of the preference sensitive conditions would you really want to have that? One of the other major things on this list was shared decision making and patient preferences, so if it's really preference sensitive care, in some ways you really want to assure you're seeing that delta in particular.

David Baker – Northwestern – Chief, General Internal Medicine Division

Yes. I agree. I mean for these preference sensitive conditions it's sure interesting to see their baseline health status as well.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. Exactly.

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

Again, so much of the longitudinal discussion I think ... measures in the short-term it may be easier to get after sort of slice in time evaluations of delta and in the long run, going after the individual data points and calculating the delta. So the difference between saying assess your functional status with a number versus assess the change in your functional status.

David Baker – Northwestern – Chief, General Internal Medicine Division

I think that's a really good point and there is a small literature on patients' ability to rate change in their health status. It doesn't always agree with some of these objective measures, but I think that's a really important thing to consider because it would sure make it a lot easier.

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

Yes.

Helen Burstin – NQF – Senior VP, Performance Measures

Okay. All right. So, moving on, the other major methods issue that was keyed up for us was specifically this issue around getting towards measures of adverse event reporting and specifically thinking about the automatic collection dissemination ... electronic systems, the capture of those events within clinical workflow and then the alert to clinicians for a possible adverse event. That's a whole lot there and I don't think, at least in looking through the Tiger Team report on patient safety there was any more specificity that I could get my hands on other than what was in those slides. So perhaps if anybody has any thoughts or has any experiences with beginning to do some of the automated adverse event reporting out of EHRs and measurements there?

Nita Roka

This is Nita. I guess I can start. We have actually— There is a project called ADESTERS. ADESTERS stands for Adverse Drug Event Spontaneous Triggered Electronic Reporting System. In that project we built ... from within two EHR systems. One is LMR, which is the EHR system of Partners Healthcare System and another product, Green Bay Medical. Some of the issues we had, as you just mentioned, like integrating the adverse event reporting into clinicians' workflow.

The second issue is the promptness and speed with which the adverse drug events are reported. Currently the physician fills out the form, which takes them three to six minutes to fill it out and fax it to the FDA, but by building this into the EHR system it takes them 60 seconds. The physician ... Partners Healthcare System; they measure the time.

The volumes of reports: In three months the FDA received 267 reports from the physician from Partners Healthcare System.

Actually, I don't know the standards and the probabilities—

(Overlapping voices.)

Helen Burstin – NQF – Senior VP, Performance Measures

I think somebody is talking on the phone. If you could go on mute that would be helpful. Thank you. I'm sorry. Keep going.

Nita Roka

Then building a clinical ... report system, I guess you mentioned that too for the adverse event reporting within the EHR system and then having a reporting system that protects the privacy of patients and clinicians. These are some of the findings from the study. Actually, it's a three-year project and the goal is to expand it to more EHR systems.

Jon White – AHRQ/HHS – Director IT

Thank you, Nita. Actually, on this particular topic I would love to hear from John Moquin and if they're still on the phone, Dan Malone and Ross, if he's come back on.

Ross Lazarus – Harvard Medical – Director of Bioinformatics, Channing Lab

Yes, I'm here. So, starting at the end, I can talk a little bit about an AHRQ project that I've had to automate vaccine adverse events, which are not quite the same. They're kind of worse than drug adverse events in the sense that we vaccinate healthy kids, so from a public health point of view it has a kind of payoff in terms of reporting, because it makes the whole vaccine project much more reliable if we know that we've got really good systems for capturing all of the adverse events after vaccination.

The project went well, but the patient died in the sense that the CDC people, who were set up to receive the notifications that we can actually automate now kind of went into radio silence mode because of the reorganization within CDC, but essentially, we showed that it's doable and it's complicated. I could talk for as long as you want, Jon, but I'll stop right there.

Helen Burstin – NQF – Senior VP, Performance Measures

Okay. Thank you.

John Moquin – AHRQ – IT Architecture Manager, Social & Scientific Systems

My experience has been less on the ... side and sort of more on the—I've been employed by Partners actually, as well as a number of other healthcare situations and vendors, so I've been mostly silent because I think what everyone has said is completely true, as I've experienced it with both issues of problem lists, issues about a unique patient identifier, sometimes even within an HIE, but certainly, in general, to not be able to catch what you might like to catch from various settings, other issues I don't think were mentioned with that word, duplicate tests, duplicate unnecessary tests. It's hard to figure out when you don't have an ability to track a patient what's already been ordered and done and where you can get information from. They really touched upon that with ... as well.

I think my caution about all of this comes from an implementation world. We talk a lot about SNOMED and we talk a lot about standard codes. From the perspective of someone who's worked in the IT department up until very, very recently, those things really haven't taken much purchase and most of the reports that you'll find for both quality and safety and other reports ... sort of customized hodgepodes, often of multiple systems that aren't terribly well integrated if they're integrated at all, which are still put together in sort of Excel or Access or, if you're really lucky, you have somebody who can put it in something like Oracle.

Hopefully this team will be able to help that change. Certainly, the carrots and the sticks in legislation will also be able to help that change, but I guess from my personal experience, the sort of current state as it is now that's it.

Helen Burstin – NQF – Senior VP, Performance Measures

I think we lost the end of that sentence.

John Moquin – AHRQ – IT Architecture Manager, Social & Scientific Systems

I said in terms of my personal experience for sort of the state of a broad healthcare market for both, large, urban, academic medical centers and small, rural hospitals, I'd say that's about it.

Nita Roka

I have a question. Here on medication safety is the forecast going to be on the drugs, not vaccines or devices, food or cosmetics?

Helen Burstin – NQF – Senior VP, Performance Measures

Again, Nita—and Tom or Cray, please weigh in here—we just included what to date the Tiger Team on Patient Safety has put forward. Medication safety was very high on their list. Anything else you want to add to that, Tom or Cray? Okay. So it looks like medication safety is fairly high on the overall priorities for that Tiger Team that they're going to present this week to the broader committee, so it just seems like it would be a good, opportune moment to think about how some of that could be built in. What are the methods issues in actually capturing some of this data?

Dan Malone – University of Arizona – Prof. of Pharm. & Public Health

Yes. It's really difficult without significant advances in coding to be able to link this information back to drugs and therefore, when we've tried to do studies of harm, looking for harm, we end up using really surrogate markers and it all gets back to the fact that there just is not a good system for coding adverse events that are drug related or possibly drug related—

Tom Sang – ONC

Excuse me. I'm sorry. We got cut off before. On your question about whether it's drugs or cosmetics or medical devices, I think we're talking about adverse drug events for now.

Helen Burstin – NQF – Senior VP, Performance Measures

Okay. Thank you.

Dan Malone – University of Arizona – Prof. of Pharm. & Public Health

So it's really important to capture this information and the current systems do a fairly poor job. It's probably because we've just done a poor job in developing the underlying architecture to code for these sorts of events. It's difficult. We've tried numerous times to pull information out of free text EHR, but it's quite challenging to be able to relate those two back. So without some significant advances I think even though it's really important it's going to be difficult.

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

With regard to doing reporting, automated reporting, I see obviously that's a very, very big quality issue. I'm not sure if it's as big of a quality measure issue. I don't see how those two fit together unless your quality measure is something as straightforward as a structural measure that you do have a quality reporting system in place.

I think from a methodological issue, if you do have access to all of the ... data I think people have been mentioning, even if you could capture sort of the universe of adverse events that actually happen, automatically capturing this and reporting it is difficult and the methodological challenge there is causality. It's very, very difficult to establish causality that A led to B. You can have sort of quality measures that get to the point of it, measures looking at diagnoses of, let's say, anaphylaxis if someone does code for anaphylaxis, but issues around causality are still going to remain. The only real way to really capture this issue is if you simply ask providers did an adverse event happen? If so, then go into the level of questions like AHRQ's common formats for error reporting.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

To sort of build on this, I mean I think we're talking about two types of problems. One is, as Nita was saying, it's a workflow problem and we need to support people in the sort of manual reporting of adverse events and there is a lot of work around culture of non-punishment, such as well as having the system set up to do it.

I think what Danny is talking about is more technical solutions around having good algorithms to reach into the data and pull out and say this is an adverse event. I think that those algorithms probably aren't as tuned or optimized as we would like them. It's not an area I know real well, but I think we're also going to run into a threshold issue here again in terms of what's the right number or even directionality for adverse event reporting, because you could argue that having the numbers go up at first is probably a good thing because we're getting at all of the unreported events now and so you'd want to sort of see it go up and then go down over time, sort of seeing what does goodness look like in this is going to be hard to get at.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

John Moquin – AHRQ – IT Architecture Manager, Social & Scientific Systems

I could speak to a couple of different types of work that I've actually done, the first of which is I did part of sort of a longitudinal study that took a look at adverse events with and without harm and the documentation, both in paper systems and EHR systems. As you might expect, there is really, unless something very, very significant happens, anything that is judged to be a relatively small amount of harm or especially no harm is rarely even documented. If you give someone 650 of Tylenol and they were supposed to get 325 it may get documented that it was over given, but it certainly wouldn't be documented anywhere as an error in the EHR. ... drawn out specifically.

Some of the other work ... is we talk about some measures that were implemented ... we took some non-intensive areas and we tried to look at the frequencies of orders. Times ... were usually held that a patient had orders for the same medications put on within two hours after a medication was initially ordered or changed outside of intensive care units and specialty care units. The assumption was that that may have been due to some sort of error. Even that was a lot more work. The juice was not worth the squeeze, if you will. We did find some things, but some things it was just a follow-up decision by a clinician or a decision after talking to other clinicians to just change it without any adverse outcomes. So I would agree that it's very, very hard to program these types of things because it's hard for people to even just see them and identify them, much less figure out how to do it programmatically.

Helen Burstin – NQF – Senior VP, Performance Measures

A question for those of you closer to this than I am, but one of the things we used to do, for example, ... is look for any medication administration of Naloxone or any of the agents that we would have potentially have used to reverse misadventures, as it were. Any thoughts about perhaps being rather than just event reporting itself, using other triggers as ways to track this?

John Moquin – AHRQ – IT Architecture Manager, Social & Scientific Systems

I think those are good as long as you get to the point where you're sort of self selecting a certain type of potential harm or type of error by just reporting those out—

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

John Moquin – AHRQ – IT Architecture Manager, Social & Scientific Systems

But as long as you have that caveat that's probably the best and most useful thing that we found to do.

David Baker – Northwestern – Chief, General Internal Medicine Division

I think this a really important exploratory area, but you can imagine setting up systems so that if, for example, on the outpatient setting, which is where I tend to concentrate, a medication is stopped like an antidepressant is stopped within ten days of initiation. Now, the physician, I should say, can put in that a medication was stopped for an adverse reaction. There is a good system set up in there, but doctors under utilize it. You can imagine sending an e-mail in situations like that and having the doctor say, "Check this. What were the reasons why this was stopped?" But I really don't think that we know how to do that with any good degree of accuracy yet. Because I think we're going to have to go beyond all of these systems that we've talked about. I think these are all good suggestions, including the natural language processing, but I just know within our system people really under utilize the tools that are available.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. Okay. Does anybody have within their system a way of capturing this already? Part of the work we had done, I know for example, the Mayo Clinic has a reportable adverse health event for 20,000 patient days and medication events with harm for 1,000 patient days as some of their measures in their EHR> Does anybody else have experience with using those kinds of measures?

Ross Lazarus – Harvard Medical – Director of Bioinformatics, Channing Lab

Yes. We have a system for capturing events within 30 days of vaccination and triaging them. For example, fever, which is almost an expected response particularly to combined vaccines can be reported, but it doesn't bother the physician. Whereas, if we see ... within 30 days for the first time in the patient's history then we bother the physician and say what do you think about this. Should we report it as a vaccine related adverse event.

In terms of the drug related adverse events, I mean I know that there's a big project going on with the many ... in the FDA. I don't know if anybody on the call is on that project, but we should probably know what those people are doing because it's directly relevant to this conversation.

Helen Burstin – NQF – Senior VP, Performance Measures

I think that's why Nita is on

Nita Roka

Yes, I am the Informatics Lead and Mark Weiner, also on the call, he is one of the co-leaders on the data

Ross Lazarus – Harvard Medical – Director of Bioinformatics, Channing Lab

Okay. Perfect.

Nita Roka

We are both on

Helen Burstin – NQF – Senior VP, Performance Measures

Anything you want to add, Mark, from your perspective on the front line?

Mark Weiner

On a case-by-case basis, mostly for research projects, we've looked for kind of ... types of activity at the wrong time in a hospitalization, so we agreed upon, for instance, the difference between a prophylactic dose of heparin, Coumadin or Lovenox and a treatment dose of the same. If you come in on a treatment dose of heparin, well, that's okay. You came in with a DVT, but if you start getting these medicines on day three of a hospitalization or later then we're more apt to think that it was an in-house DVT.

We've also done some more things with looking for LFT abnormalities and following the initiation of a medication. Actually, we're coming out with a paper about category X medications given to people, who are known to be pregnant, but of course, the conclusion is that we weren't able to detect the pregnancy in terms that were available in a decision support way to prevent the category X medicine and then most of the people who turned out to be pregnant on a category X medicine were on it, on the medication previously and then had a pregnancy afterwards and stayed on their medicine a little too long.

Helen Burstin – NQF – Senior VP, Performance Measures

Very helpful. Anything else on this broad area of adverse event reporting in EHRs as being some methods issues that would need some work? I agree with David. It sounds quite exploratory, more so than some of the areas we've talked about before.

David Baker – Northwestern – Chief, General Internal Medicine Division

But again, do you have a specific goal in mind where we're looking for adverse events that are sort of idiosyncratic reactions or adverse events that are active, but regrettable misadventures or drug interaction types or really, all of the above?

Helen Burstin – NQF – Senior VP, Performance Measures

I don't think we have a good handle on that yet. Tom or Cray, anything you want to add there in terms of exactly what they've been looking for on the Med Safety Group?

Tom Sang – ONC

No, but it rings a bell.

Helen Burstin – NQF – Senior VP, Performance Measures

Nicely done. Okay.

Cray Noltert – ONC

Can you narrow down the question?

Helen Burstin – NQF – Senior VP, Performance Measures

Sure.

Cray Noltert – ONC

... the domain or sub-domain?

Helen Burstin – NQF – Senior VP, Performance Measures

Right. So the specific question was this is obviously a pretty big area and at least just what I could see on what had come out of the Patient Safety Tiger Team was adverse drug event reporting. Was there any greater degree of specificity of the kinds of reactions they were looking for or the kinds of events they were looking for?

Leah Marcotte – ONC

I was on the ONC ... Patient Safety Tiger Team. We did talk about targeting specific high impact reactions or medications, but no one came to kind of a specific example of what those might be.

M

So, for example, Leah had found a measure the MD ... the Medicare—

Leah Marcotte – ONC

... medication safety, like MPMS—

M

The MPMS, which is a Medicare measure for very, very specific drugs. It's heparin and warfarin adverse events and exposure to hypoglycemic drugs, but there are very definite limitations to that measure. It's mostly inpatient. It's only Medicare beneficiaries and it's all claims based data. So I guess first would be to see if there are any groups out there looking at ambulatory or a wider spread measure.

Number two: If there are methodologic solutions in terms of looking at preventable adverse event reporting out of the EHR system. I guess those would be the two main issues. The other issue is that the adverse events or ... events right now is a voluntary phenomenon going to the FDA system. Is there a process to automate this? Those are some of the, I think, issues that the group is dealing with. So, a lot of it's really exploratory.

Helen Burstin – NQF – Senior VP, Performance Measures

Got it. Okay. I mean there certainly are some measures NCQA and others have looked at about what's the most important drugs to have monitoring annually for persistent meds, drug ... in the elderly, some pregnancy related measures and some Warfarin INR monitoring measures that are probably those high impact meds that you want to be sure you capture, but it's a little different than the automated reporting piece.

Dan Malone – University of Arizona – Prof. of Pharm. & Public Health

I just wanted to add in a real quick comment about drug interactions. There is another committee or another scope of work that's trying to identify a small sub-set of drug interactions that clinicians should be alerted on, so that working group is headed up by Brigham Women's researchers.

In addition, the Pharmacy Quality Alliance has developed and is testing a set of drug interaction pairs that would hold pharmacies accountable on a quality measure. So there is some work being done on the drug interaction space. It kind of relates to EHR, but more on the e-prescribing side. That's something that's happening concurrently.

Helen Burstin – NQF – Senior VP, Performance Measures

That's great. That's some important work. Has there been any additional work? Would some of this try to sort of narrow the filter of the over alerting that a lot of clinicians get ... when I use my EHR and so I ignore everything—

Dan Malone – University of Arizona – Prof. of Pharm. & Public Health

Yes. I think that's one of the goals is to figure out which of these drug pairs should you receive an alert on, because right now our anecdotal experience is that when they turn the systems on, because of the over alerting issue, they get turned off fairly quickly or just ignored.

Helen Burstin – NQF – Senior VP, Performance Measures

Or just ignored. Yes.

Dan Malone – University of Arizona – Prof. of Pharm. & Public Health

We've done some work at the VA. Even though they've tailored their system they still get a lot of ASBF in required fields. If the F is your left hand on the keyboard you get a lot of sensical responses to acquired acknowledge of an interaction alert along with other types of alerts. So they're trying to narrow down which ones should we really be trying to bring to clinicians' attention at the time they're prescribing and then also we're doing a bunch of work within pharmacies about which ones should pharmacies see—

Helen Burstin – NQF – Senior VP, Performance Measures

Right.

Dan Malone – University of Arizona – Prof. of Pharm. & Public Health

I guess it goes beyond that from the standpoint of for most of these interactions there's maybe a theoretical basis for it, but the clinical evidence behind many of these is lacking. That's a whole other issue, but lots of stuff happening in that space.

Helen Burstin – NQF – Senior VP, Performance Measures

Very good. Exciting. All right. I think we may have finished that particular method issue unless anybody else has anything else they want to add in here. Great. You guys are being efficient.

I think it would still be useful, if people are willing to do it, to go back to that Tiger Team summary perhaps and just walk through the other sub-domains and measure concepts that are listed; again, they're quite broad; and see if there are any other methodologic issues that we'd want to highlight as some of these other ideas come forward. Does that sound reasonable?

W

Yes.

Helen Burstin – NQF – Senior VP, Performance Measures

Good. Okay. So let's start at the top, the general issues of care coordination: We've talked a lot about some of the key need to be able to do the longitudinal follow-up to get at things like follow-up labs, med req, follow-up after transition. Anything on that top list of those four sub-domains that bring up additional items, additional methodologic issues that would be important to consider? Okay. I think we've captured much of that.

The one thing we haven't really spent a lot of time on, but I think it's being captured by the Patient and Family Engagement Group is how do we handle the patient reported information, for example, on lack of follow-up, etc. or the transition experience. We could probably come back to that.

M

The one I had a question on, Helen, was can you say more about diagnostic management?

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. That group is still sort of working through those. I think the issue was trying to ensure that the patient with a given condition gets the right kind of interventions at the right times. So, for a patient with a given condition are you going through the logical phases of a depression you have to ... your chronic management is a patient with a given condition getting the meds they should get. So it's a combination, I think, of effectiveness, as well as sort of timely follow-up and coordination.

M

Sort of adequacy and appropriateness of testing after a new diagnosis?

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

M

Or testing and treatment after a new diagnosis. Okay. Thank you.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. All right. How about efficiency? ... is obviously a difficult one, but perhaps it's easier to capture with interoperability across institutions. Other ones here: Diagnostic imaging. Appropriate medication use. We spent a lot of time on the idea of a longitudinal dashboard for leading conditions. Anything else there people want to highlight? All right.

Patient safety—

Bob Dolin – HL7 – Chair Elect

Can I throw in one?

Helen Burstin – NQF – Senior VP, Performance Measures

Please do. Yes.

Bob Dolin – HL7 – Chair Elect

For the efficiency, this is something that we've been struggling with. If we start thinking about over use measures, the over arching problem I think for us is that we're not capturing the results of some of these tests in queriable fields.

An example: One of our fellows is working on a project right now looking at over use of pap smears. We just wish it was a lot easier to be able to get that data out, but it's all driven by the longitudinal data is multiple pap smears over the preceding; somebody has had three normals in the last three or four years, but then it's also if they've had an HPV test that's negative even after one year that would qualify them for three year follow-up.

We've talked about this as well on the appropriateness of the rate of normal cardiac caths or normal stress tests as being a potential indicator of over use. So that's kind of a cross cutting issue that if we don't have this data in queriable fields we're not going to be able to look at efficiency well.

Helen Burstin – NQF – Senior VP, Performance Measures

Great. Did that stimulate anything else for anybody else on the call?

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

Yes. I mean I think we're talking about two different kinds of things. One is did they need it at all type of measures. Is this either the right test or should they have had any imaging, like in back pain or the too frequent breast cancer screening or cervical. Then there's the repeat imaging or repeat testing, because the original tests are not available to me. That may be solved by the exchange. I think we're going to need to think about the impact of missing data on this, so can we really ding one clinician if the other one, who did the test on the patient, didn't upload it or share it, which I guess is as much a policy question as a methods question, but it does get back to missing data.

Bob Dolin – HL7 – Chair Elect

Well, maybe that's not a quality measure on which to ding the provider, but it is something that in the future of meaningful use where there will be interoperability and more availability of prior testing that we ought to be able to have a baseline level of testing now and see improvement later, sort of a benefit of meaningful use as opposed to some provider quality.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

Or the individual clinician's measure is what proportion of your labs, imaging, whatever, is uploaded to the interchange and is uploaded using structured reporting so that it can be grabbed again rather than just as text blob or an image.

Helen Burstin – NQF – Senior VP, Performance Measures

Text blob, I like that. Okay. Very good. Going down the list there, where were we? We were on imaging. So looking at patient safety there, I think we've talked about ADEs. We've talked about a lot of the issues around med req. Any thoughts about some of these other patient safety areas that are identified? Hospital associated events, such as Hays or HEE as an example or being able to, for example, do a risk assessment and a rate of an outcome, the falls example here.

I'll throw out one for an example. The HII data is very frequently if you're looking at somebody whose surgical site infection for example where somebody has an implanted device you're supposed to go out a full year, for example, to look at SSI for that population. Obviously, pretty difficult; some real difficult issues there in terms of somebody may have flown somewhere else to get a specialized procedure done and then go back home and no way of kind of keeping track of that.

Abel Kho – Northwestern – Assistant Professor, General Internal Medicine

Yes. I'd reflect that. I think ... is a really difficult space. Surgical site infections are a great example of that. Even looking at some definitions around healthcare associated MRSA, for example, also pretty complicated, so that one might be one we might want to push until later.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

I think we'll have to think about it in terms of policies about not paying for never events or for defects and so are hospitals likely to put in place a system to capture these sorts of things in the context where they won't get paid for it if they find it?

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. Fortunately, that's out of our purview.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

Yes. Yes. No.

Joanne Cunniff – AMA-PCPI – Director

Again, you may have discussed this in past calls, but is there a way to capture a patient that has received the full evidenced-based standard of care for VTE prevention and still has a VTE?

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

Joanne Cuny – AMA-PCPI – Director

I suppose the same could be true for hospital acquired infections, but I think in VTE it's very straightforward what is the evidence based standard of care. Then again, patients often will still have VTE.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. VTE is actually a somewhat easier one just because I think at least the process measures and the outcomes there are pretty well set—

Joanne Cuny – AMA-PCPI – Director

Yes.

Helen Burstin – NQF – Senior VP, Performance Measures

But again, I think it raises the same issue of somebody may go home after their total knee replacement and be very disconnected from the place they had their total knee done and the actual ... of the Coumadin or whatever may not proceed and they may get a VT. How do you track that?

Joanne Cuny – AMA-PCPI – Director

That's right. The other issue around that is that in some cases a patient maybe would be leaving the hospital and although some sort of anti-thrombotic or anti-platelet agent may be prescribed, but it won't be initiated until a day or two after the patient is out of the hospital and sometimes, for instance, Warfarin can't be given the same day the patient leaves the hospital, but in fact, the drug is to be filled. It doesn't appear as a drug that the patient was receiving the last day in the hospital and so it looks like it wasn't prescribed. As I say, you may have discussed this already, but I do recall this as an issue in some of the settings where we've studied.

Helen Burstin – NQF – Senior VP, Performance Measures

Okay. Very good. All right. Patient and Family Engagement. We don't have to do too much here since the Tiger Team has taken care of some of the health status and function status issues. Any thoughts about the methods issues involved in, for example, incorporation of patient preferences or, for example, if a patient goes through a shared decision making tool, being able to capture some of that? Anybody have any experience there?

David Baker – Northwestern – Chief, General Internal Medicine Division

We don't have any experience yet, but the head of patient education here is just starting to look at some of these commercial tools that can be integrated in with the electronic health record and a patient viewing some of these or using clinical decision support. In some of them, at least, they say it will be automatically captured. That's all we know about that and we're looking into that more, so I think that's coming. It doesn't mean still that they understood the information, but I think that that's probably going to be as close as you can get to a process measure. If they went through an established tool or system that's been shown to be the best possible then that certainly is more meaningful to me than they were just handed something at discharge.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. That's interesting. The Foundation for Medication Decision Making has some emerging issues of decision quality where that also includes a patient knowledge, pre and post-assessment, which has turned out to be fairly powerful as well, but again, another very difficult thing to think about incorporating into an EHR space.

David Baker – Northwestern – Chief, General Internal Medicine Division

Yes. That's interesting.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. Anybody else on the Patient and Family Engagement piece, either preferences, health outcomes, self management? Ross, if you're still with us, there was one specifically here looking at being able to understand health activities, coordination and connection to community resources. Any thoughts on that one? He was going to pop in and out; perhaps he popped out. All right. We'll move back to him on that.

Lastly here, I wish he was still with us, is the population health/public health domains here. Any thoughts? I think some of the longitudinal ... issues we've already talked about at the front end here. Specific thoughts about some of the sub-domains they picked around smoking, BMI, alcohol and healthy behaviors or blood pressure, diabetes, mental health on the preventive outcome side there?

David Baker – Northwestern – Chief, General Internal Medicine Division

Helen, maybe one for mental health. I get the sense that primary care clinicians and psychiatrists use mental health diagnostic terms to mean very different things. In this new world order it used to be that people would do coding after the fact, look at all of the chart documentation and then based on their complete knowledge of an ICD code they would pick the best ICD code so you'd have some internal consistency. But now what we see is people are selecting something on a problem list based on their interpretation of a display name and all of the coding is driven directly from that, so a primary care doctor may say that you're depressed or a psychiatrist may say you're depressed and they don't necessarily meet the same criteria. So there is probably some issue there with knowing who the provider is and what you think their interpretation is of a particular diagnosis.

Helen Burstin – NQF – Senior VP, Performance Measures

Interesting. We've also been doing some work recently with a set of measures going through our process that look at the delta of the PHQ-9 using something more objective and patient based in terms of depression symptomatology. Again, David, that's one of those other conditions that gets added to that list of the sort of pre-post list we had earlier we were bantering around; low back and hip and knee, etc.

Daniel Rosenthal – National Quality Forum – Senior Advisor, HIT

For the health equity sub-domain I think that some health equity issues can be identified by using a traditional, clinical measure and simply stratifying it based on, let's say, age or gender or whatever health disparity cross section you want to look at. So a methodological issue there will be our current measures, if not currently specified, to address a specific type of stratification is that acceptable use of the measure. In a diabetes measure if it's specified for all ethnicities can you use the same measure and report based on a stratified base on ethnicity. Has it been tested and validated for that particular use?

Helen Burstin – NQF – Senior VP, Performance Measures

Okay. Anyone else on the equity issue? All right. I'm getting the sense there's a decreasing energy. You guys have worked very hard. Any other methods issues that come up that we haven't talked about that would be worth exploring as long as we have strong people, like you on the phone? I guess the one question I would think of; sorry to pick on you, David Baker; would be this question of given your expertise whether there are some important issues here around health literacy if we want to start getting the patient to enter some of these data. Any thoughts on that as a methods issue?

David Baker – Northwestern – Chief, General Internal Medicine Division

I think, both for the patient and the physician that is one of the challenges; getting accurate data in. Beth Hahn is a researcher, who has done some really nice work with what she calls the talking touch screen. She showed that basically it's touch screen with headphones and the questions are read and the options light up and she's shown that even low literate patients can complete very well a lot of these health status questionnaires that have been used in some other things, so I think the technology is getting to the point where they should be able to overcome a lot of the literacy barriers.

Helen Burstin – NQF – Senior VP, Performance Measures

Great. I've also been impressed that the nursing assistants in our clinic are administering the PHQ-9 in Spanish and just entering it right into the EHR in just minutes, so another potential way. It's still the voice of the patient, but it's read to them in a fairly simple way.

Bob Dolin – HL7 – Chair Elect

Yes, we've started doing the PHQ-9 by the way. I agree with what you said before; that that should be added to the list. It's really been great. The docs have like it and I certainly think it's improved my care. I don't think it takes additional time. That's the issue for a lot of physicians; there is some intrinsic resistance to going to questionnaires. I will say definitely the first few times you do it it feels a little bit awkward—

Helen Burstin – NQF – Senior VP, Performance Measures

Yes.

Bob Dolin – HL7 – Chair Elect

Rather than doing an open ended interview to actually doing this questionnaire and pushing it off to room nurses and others is probably not something that we want to advocate, because I'll tell you, I've had patients break down crying when they get to question four and that sure tells me a lot more than just their score.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. Good point. Okay. All right. Tom, Cray, Jon, anything else we should be concentrating on? I think we were somewhat efficient. I assume we need to do public comment as well, yes?

Jon White – AHRQ/HHS – Director IT

We do. We need to open it up to five minutes for public comment or if there are any comments at all.

Phil Renner – Kaiser Permanente – Principal Consultant, Metrics Development

Maybe before we do that there's actually one other thing that I wanted to add that was a little bit outside of all of this. We've got to test these things and we've got to test them in the context of the measurement program. I mean I think one of the things we struggled with with the phase one measures and taking measures that may or may not and generally not testing within an EHR setting, putting them into the QDS, which was a reference data model, but again, one that the mapping hadn't been tested and then the measures put into the EQMF, that's all I think being implemented without being tested. So we need to make sure that not only are the measures themselves tested, but in this particular setting and with real life workflows not sort of idealized workflows so that we can know what we're getting.

Helen Burstin – NQF – Senior VP, Performance Measures

Yes. Good point. Okay. Operator, do you want to see if there is anybody, who would like to make a comment?

Jon White – AHRQ/HHS – Director IT

While we're waiting for public comments to queue up I want to, on behalf of my colleagues at ONC and for myself also, express my extraordinary gratitude to the wonderful folks that are on the phone. It's one thing to ask all of the world's experts to participate in a process like this and it's another thing to ask them to do it with like a week's notice, so we really appreciate both your deep expertise and your willingness to make yourself available on really short notice to do this. Thank you so much.

Tom Sang – ONC

I definitely second Jon's comments. I appreciate all of the hard work and the tremendous effort that you guys put in in such a short time, so thank you a lot.

Operator

We have a public comment.

Carol Bickford – American Nurses Association

This is Carol Bickford from the American Nurses Association. Your listing of the issues in relation to the methodologies of capturing the quality indicators and the adverse outcomes and those sorts of things was focusing very much on the physician's perspective. Would you speak to the identification of measures that go across the full spectrum and what do you envision would be the complexities associated with an intra-professional problem list or some other quality measures? Do you think that it would be exponentially more complicated or would it facilitate the quality of care measures?

Helen Burstin – NQF – Senior VP, Performance Measures

I think we were trying not to be quite so physician centric. I think the problem list in many places is something that flows across multiple sites and specific clinicians. I often find some of the social issues that I don't tend to necessarily put on my problem list as something my nurse will pop on or the nurse practitioner or the PA, so I think there is a real hope that that inter-disciplinary approach to ensuring that the right things are put in front of the clinician at the right time is really the way we'll drive improvement. Any thoughts from the group here?

Dan Malone – University of Arizona – Prof. of Pharm. & Public Health

I would comment that it also cuts across the pharmacy spectrum as well and from the pharmacy side there is a lot of activity within the HIT realm and software development and various vendors, including those problem lists. So it does cut across care, points of care and provider types.

Operator

There are no other comments at this moment.

Helen Burstin – NQF – Senior VP, Performance Measures

Jon or Cray or Tom, want to talk about next steps here?

Jon White – AHRQ/HHS – Director IT

I'll briefly throw those out. After having captured today's conversation I think what we're going to do is look to turn around a listing of the methodologic issues that have been identified with all of the different domains of quality measures and try to get that in shape to be able to be presented Thursday. Although aspirationally I would like to be able to share it with the Tiger Team to make sure that we got it right, I'm not sure that we're going to be able to promise to do that. If we can, you will see it come across your e-mail.

Then, on Thursday all of the different Tiger Teams on the Quality Measures Workgroup are presenting at a public hearing. Helen and I will be presenting the work of the group to the Quality Measures Workgroup. So that's the next steps in the process.

Tom or Cray, have you got anything to add?

Tom Sang – ONC

No, Jon. I think you've summarized it pretty accurately. Thank you.

Jon White – AHRQ/HHS – Director IT

Excellent.

Helen Burstin – NQF – Senior VP, Performance Measures

Great. We can let everybody go 45 minutes early.

Jon White – AHRQ/HHS – Director IT

We appreciate your time. Thank you, Helen, so much for also being willing to lead on short notice. As always, it is deeply appreciated.

Helen Burstin – NQF – Senior VP, Performance Measures

My pleasure. I guess we'll be back in touch with this group, perhaps to take a deeper dive at a later date, yes?

Jon White – AHRQ/HHS – Director IT

That is a high potential, so watch your e-mails.

Helen Burstin – NQF – Senior VP, Performance Measures

All right. Thank you, everyone.